A survey-lite of tracking-based soccer research

SMGT 432

October 30, 2023

Outline

2016:

- Spearman (Opta)

2017:

- Spearman (Sloan)
- Power *et al.* (KDD)

2018:

- Spearman (Sloan)
- Fernandez and Bornn (Sloan)

2019:

- Fernandez *et al.* (Sloan)
- Shaw and Glickman (Barcelona)

2021:

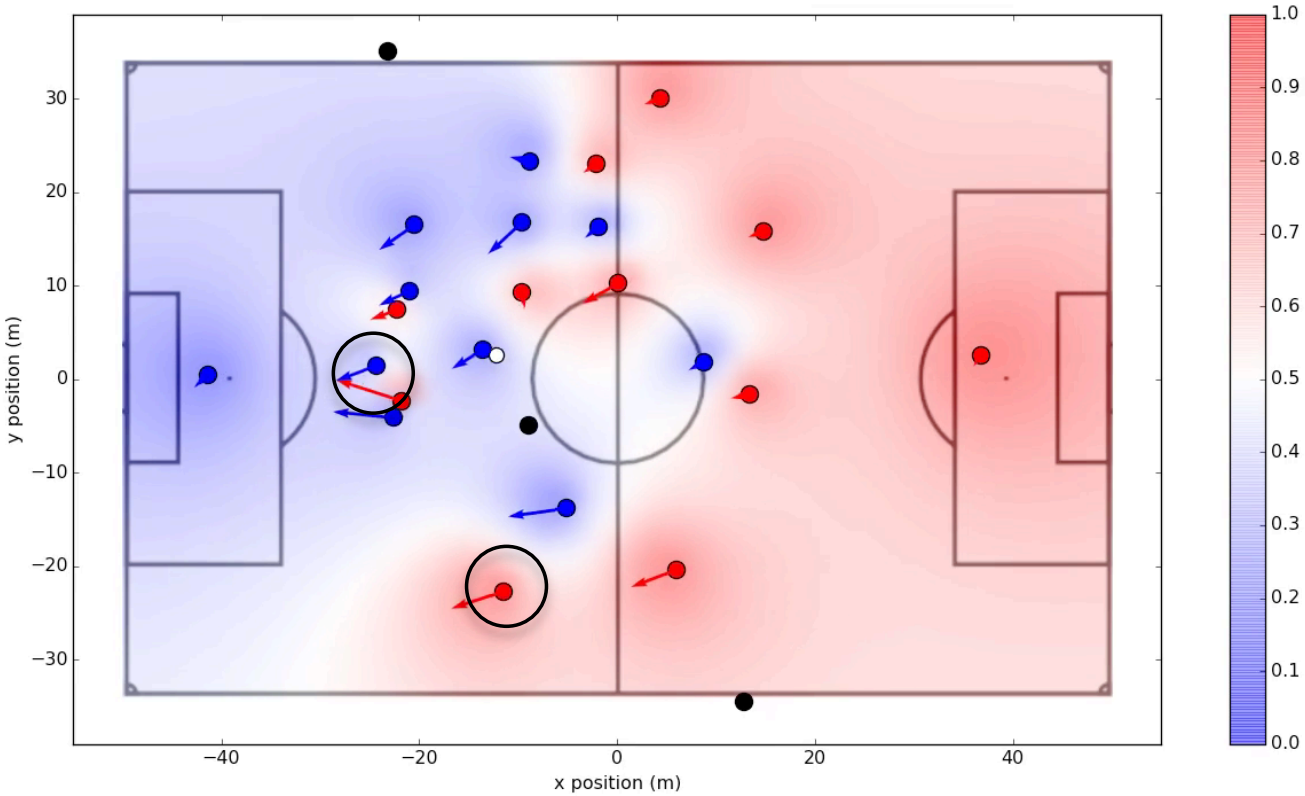
- Shaw and Gopaladesikan (Sloan)

2022:

- Everett *et al.* (StatsBomb)

Spearman (2016 Opta Forum)

Quantifying Pitch Control



$$PCF(t_i, \ell_i) = \left[\frac{\sum_i \ell_i t_i^{\beta}}{\sum_i t_i^{\beta}} + 1 \right] /2$$

Spearman (2016 Opta Forum)

Quantifying Pitch Control

Data:

- TRACAB (provided by the forum)

Calculating times using:

- Player position
- Player velocity
- Player acceleration
- Maximum player speed

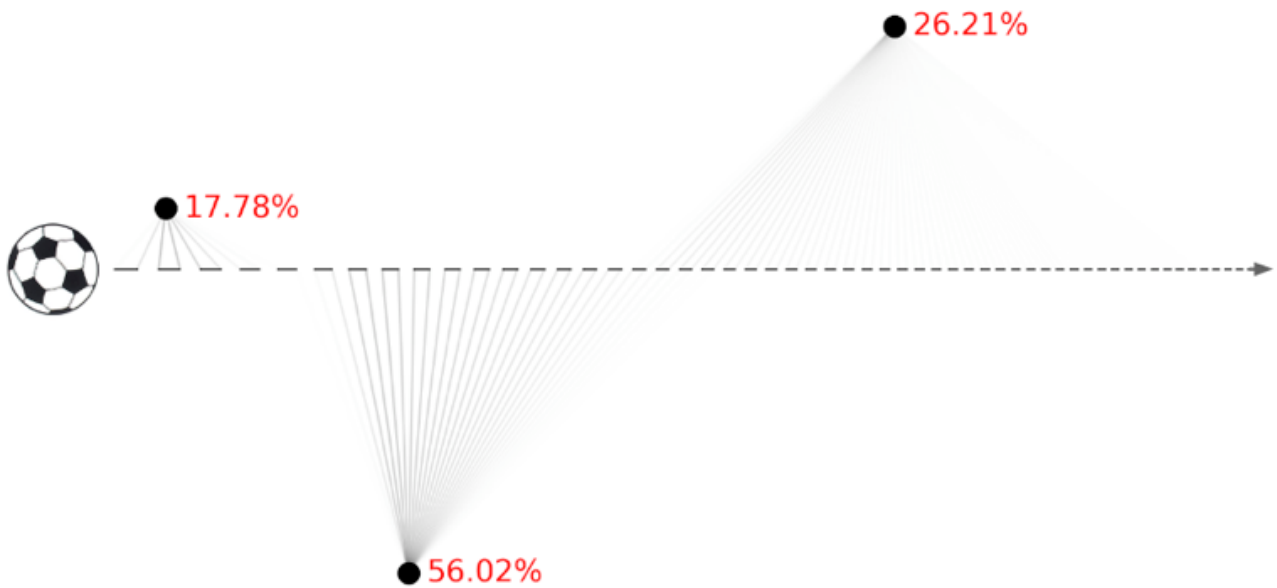
Applications:

- A new way to watch film
- A new metric for player performance
- Player positioning

Citations: 7

Spearman *et al.* (2017 Sloan Conference)

Physics-Based Modeling of Pass Probabilities in Soccer



Spearman *et al.* (2017 Sloan Conference)

Physics-Based Modeling of Pass Probabilities in Soccer

Data:

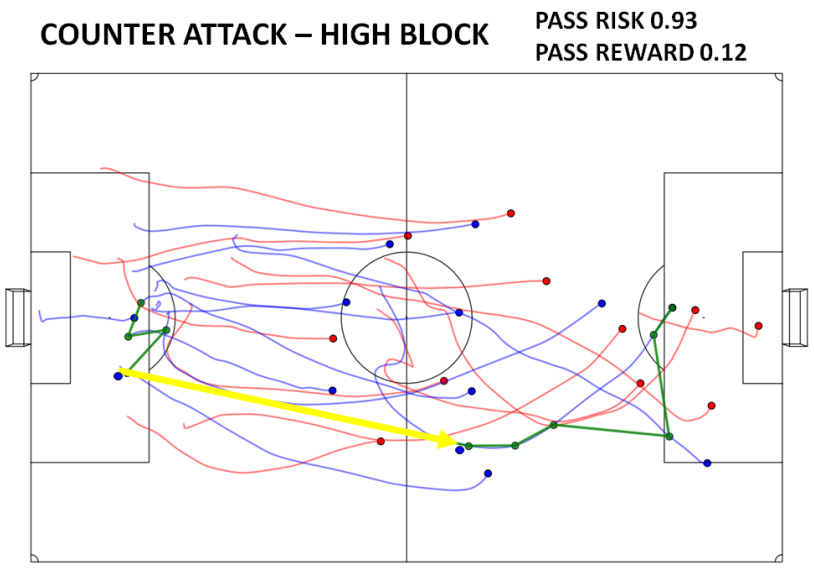
- 38 matches from 2015-16 EPL (provided by Crystal Palace)

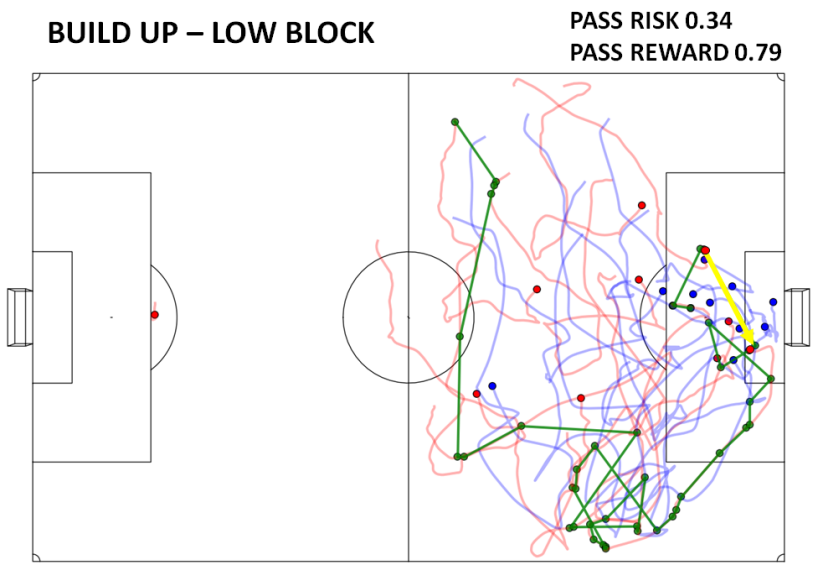
Applications:

- Pitch control
- Hypothetical passing

Citations: 101

Power *et al.* (2017 KDD Workshop)





Power *et al.* (2017 KDD Workshop)

Data:

- 726 matches from 2014-2016 EPL (provided by STATS)

Features:

- Speed of the player in possession and the intended receiver
- Speed of the nearest defender toward the passer and the receiver
- Distance of nearest defender to the passer and receiver
- Nearest defender angle to the passing line
- First time pass
- Time from regaining possession

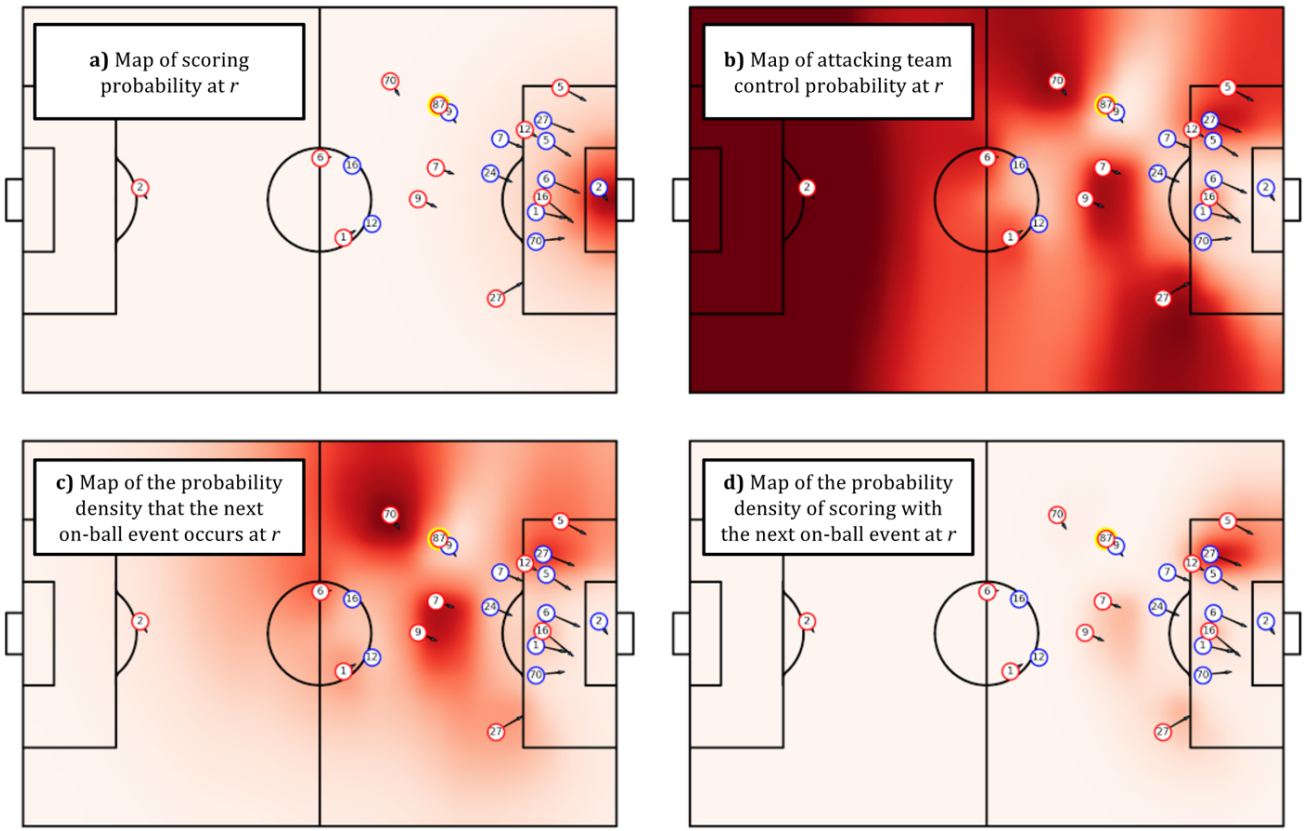
Applications:

- Match analysis
- Ranking the riskiest players
- Ranking of best players receiving passes

Citations: 133

Spearman (2018 Sloan Conference)

Beyond Expected Goals



$$P(G|D) = \sum_{r \in \mathbb{R} \times \mathbb{R}} P(S_r|C_r, T_r, D)P(C_r|T_r, D)P(T_r|D)$$

Spearman (2018 Sloan Conference)

Beyond Expected Goals

Data:

- 58 matches of tracking from 2017-2018 (provided by Hudl)

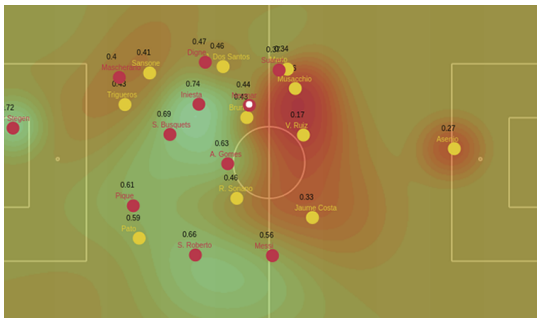
Applications:

- Tactical moment analysis
- Match analysis
- Team performance
- Player performance

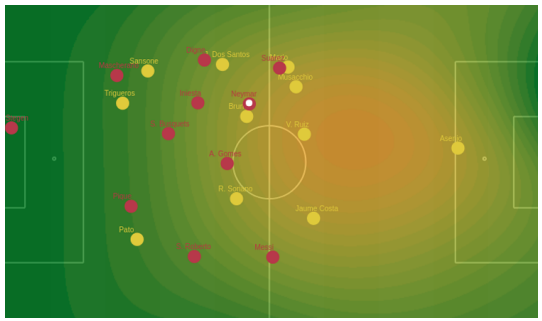
Citations: 133

Fernandez and Bornn (2018 Sloan Conference)

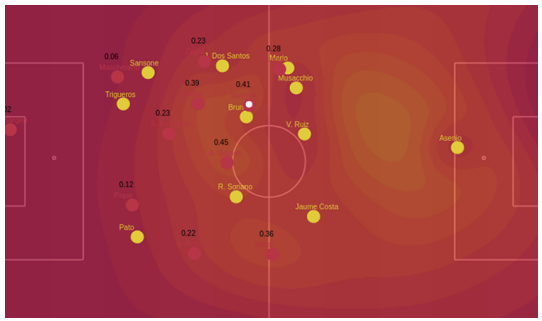
Wide Open Spaces:
A statistical technique for measuring space creation in professional soccer



(a) Pitch control surface



(b) Pitch value based on ball position



(c) Value of the owned space as product of pitch control and field value

$$Q_i(t) = PC_i(t)V(t)$$

Fernandez and Bornn (2018 Sloan Conference)

Wide Open Spaces:
A statistical technique for measuring space creation in professional soccer

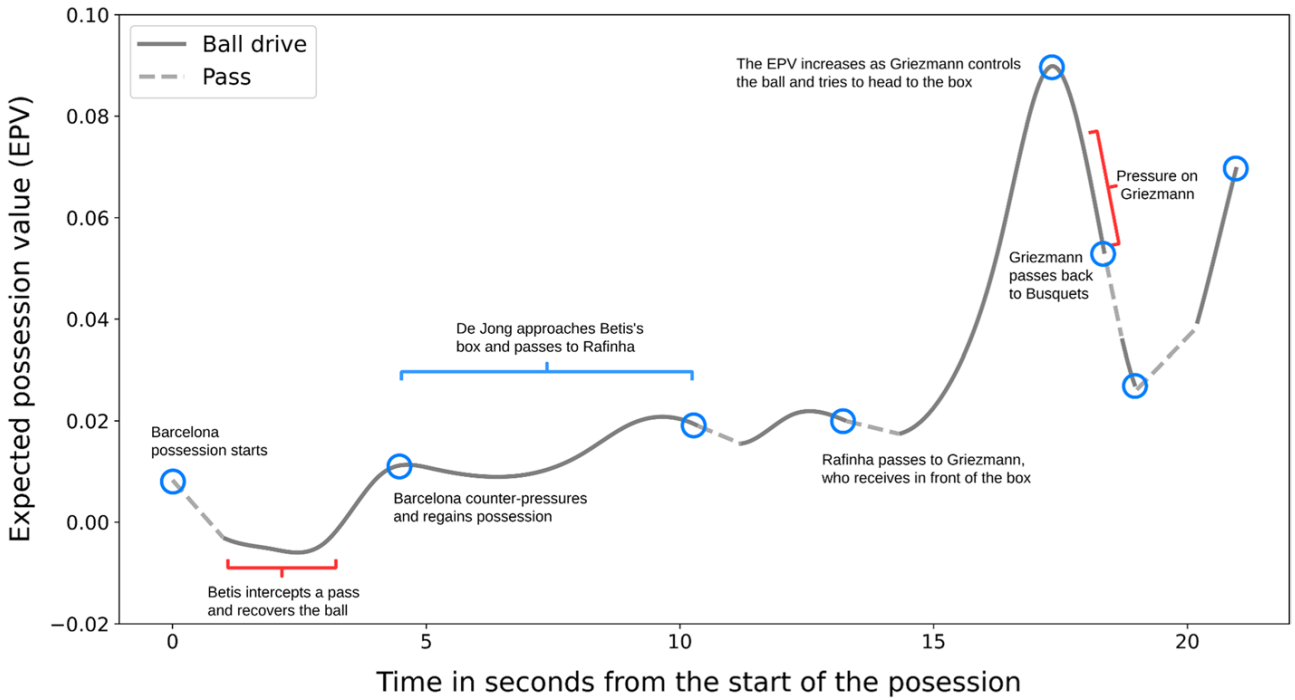
Data:

- 20 matches of Metrica from Spain (provided by Barcelona)

Citations: 170

Fernandez *et al.* (2019 Sloan Conference)

Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer



$$EPV(t) = E[X|A = \rho]P(A = \rho) + E[X|A = \varsigma]P(A = \varsigma) + E[X|A = \delta]P(A = \delta)$$

Fernandez *et al.* (2019 Sloan Conference)

Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer

Data:

- Tracking data from 2012-13 EPL (provided by STATS)
- Footovision from 2017-18 and 2018-19 FC Barcelona matches (provided by FC Barcelona)

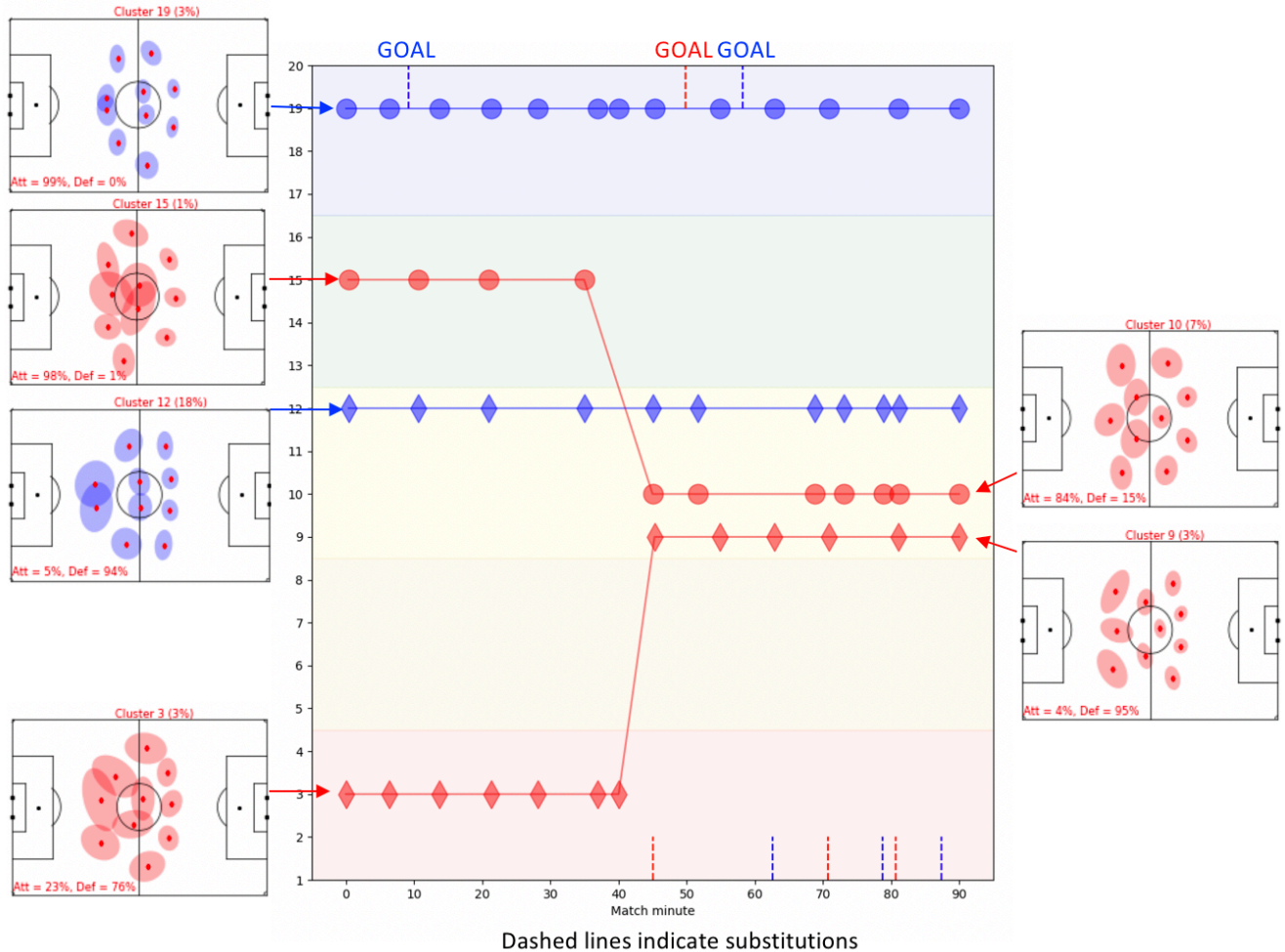
Applications:

- Pass analysis
- Distilling off-ball value creation
- Decision-making analysis

Citations: 148

Shaw and Glickman (2019 Barcelona Summit)

Dynamic analysis of team strategy in professional football



Shaw and Glickman (2019 Barcelona Summit)

Dynamic analysis of team strategy in professional football

Data:

- 180 matches from "an elite professional league"

Applications:

- Exploit opposition tactical changes
- Identify weaknesses of specific formations
- Consider formation in specific phases of possession

Citations: 37

Shaw and Gopaladesikan (2021 Sloan Conference)

Routine Inspection: A playbook for corner kicks

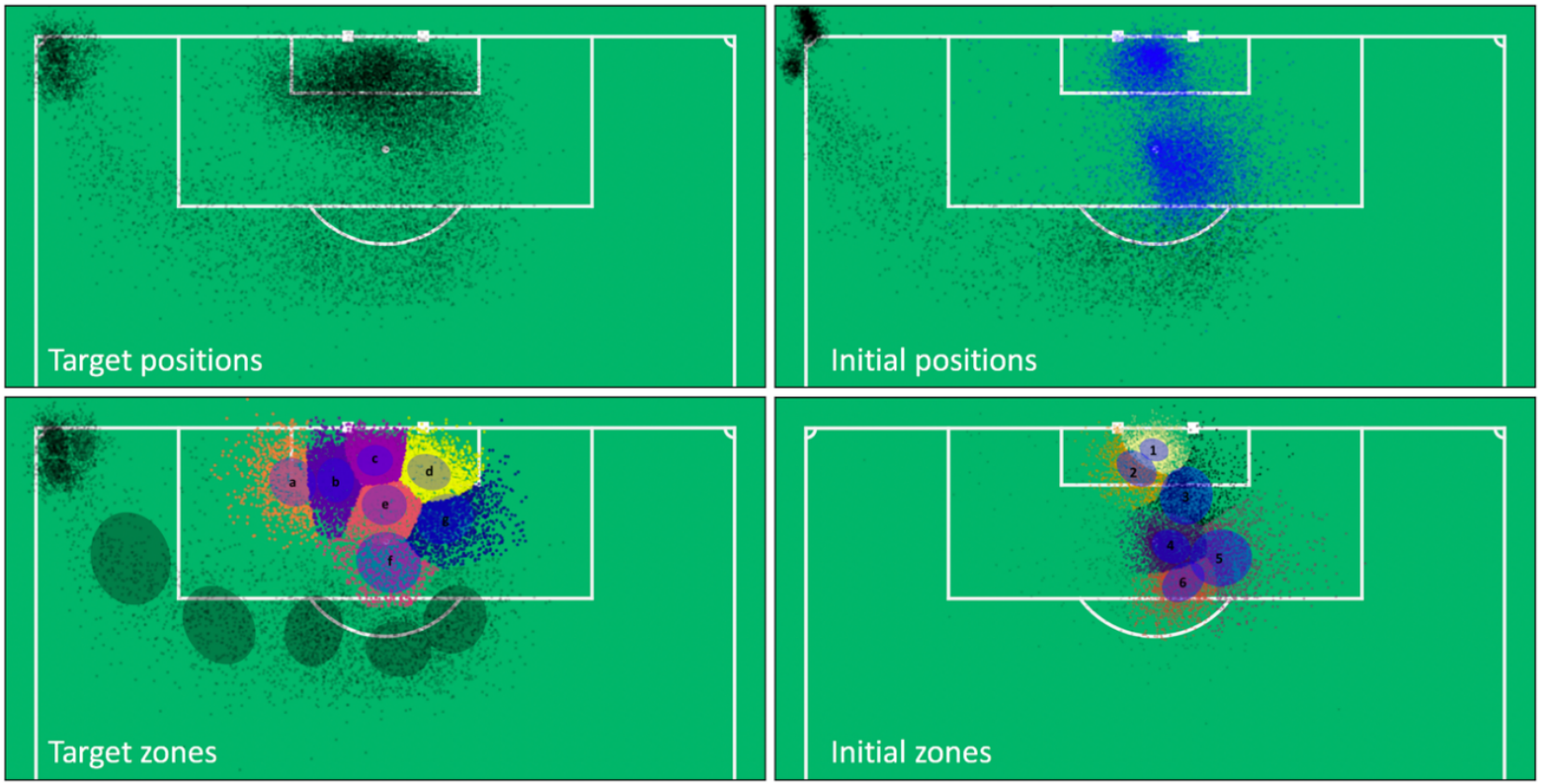


Figure 2: (*upper left plot*) The target positions of all ~15000 attacking players in our sample of corners. (*lower left*) The results of the 15-component GMM fit to the target positions - the seven 'active zones' in the penalty area are represented by blue ellipses and labelled *a-g*. (*upper right*) The initial positions of all 15000 attacking players - players colored blue are tagged as 'active'. (*lower right*) The results of a 6-component GMM fit to the initial positions of the active players.

Shaw and Gopaladesikan (2021 Sloan Conference)

Routine Inspection: A playbook for corner kicks

Data:

- 234 matches from "an elite European professional league" (presumably provided by SL Benfica)

Methods:

- Gaussian mixture modeling
- Non-negative matrix factorization
- Gradient boosting (for defensive role classification)

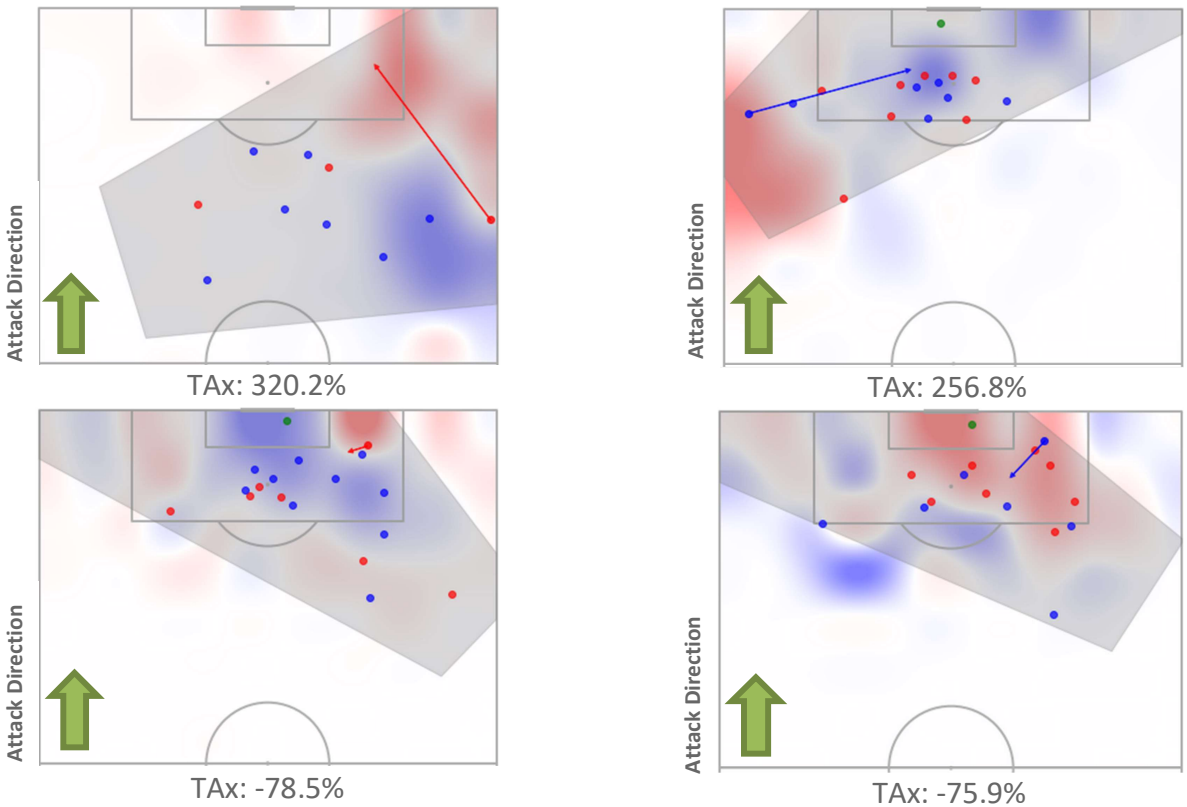
Applications:

- Analysis of an opponent's offensive corner strategies
- Comparing the effectiveness of zonal systems
- Training optimization

Citations: 12

Everett *et al.* (2022)

Contextual Expected Threat using Spatial Event Data



$$\text{TAx} = 100 \times \frac{\text{xT}_{spatial} - \text{xT}_{original}}{\text{xT}_{original}}$$

Everett *et al.* (2022)

Contextual Expected Threat using Spatial Event Data

Data:

- SB360 data from 2021-22 EPL (provided by StatsBomb)

Methods:

- Convolutional neural network

Applications:

- More accurate possession value model
- Thread Above Expected
- Defensive Optimizer

Citations: 1